Solution Error Modelling, Inverse Problems, and Uncertainty Quantification

Mike Christie

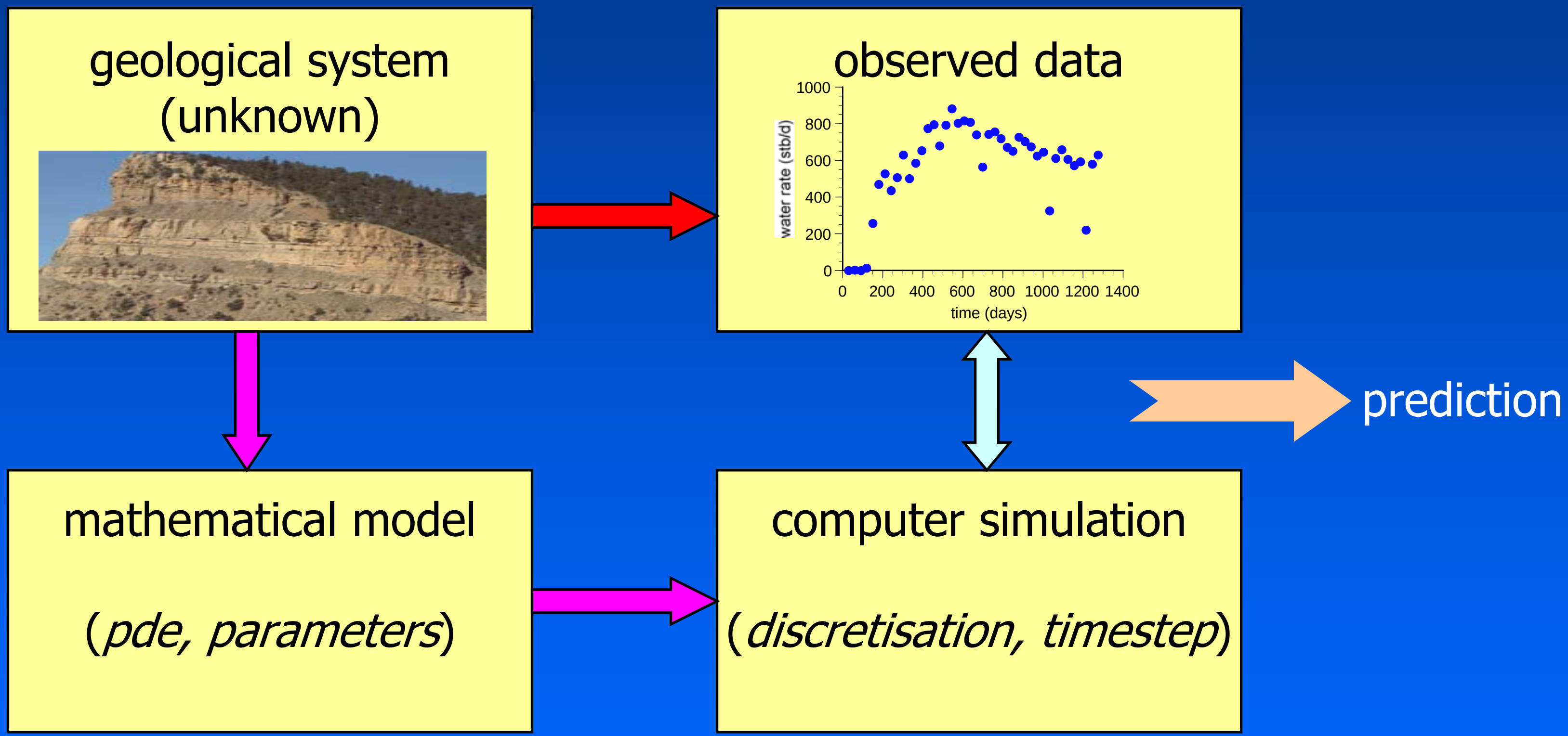
Heriot-Watt University

Co-workers:

Alannah O'Sullivan, Heriot-Watt

Dave Sharp, Los Alamos National Laboratory

Introduction



How to Determine Likelihood

- Likelihood is probability of difference between observation and model

$$obs - sim = (obs - true) - (sim - true)$$

$$= err_{obs} - err_{sim}$$

- Likelihood is the probability that the observed error and the model/simulation error add (subtract) to give the discrepancy

Detail in Models

- Many approaches use simple models to get approximate match and then add detail
- Examples
  - Increased physics (eg coupling particles with continuum approaches, heat flux in ICF)
  - Mesh resolution (coarse meshes may damp out physics that needs certain level of resolution)

Simulation Errors

- Assume a model

$$P = f_M\left(p_1,\ldots,p_n,\theta_1,\ldots\theta_k\right)$$

  - $p_1,\ldots,p_n$ are parameters that are varied
  - $\theta_1,\ldots\theta_k$ are unknown initial conditions or parameters describing additional physics (or sub-grid physics)

- Simulator prediction

$$P_S = f_S\left(p_1,\ldots,p_n,\theta_1,\ldots\theta_k,\Delta x,\Delta t,\ldots\right)$$

- Error

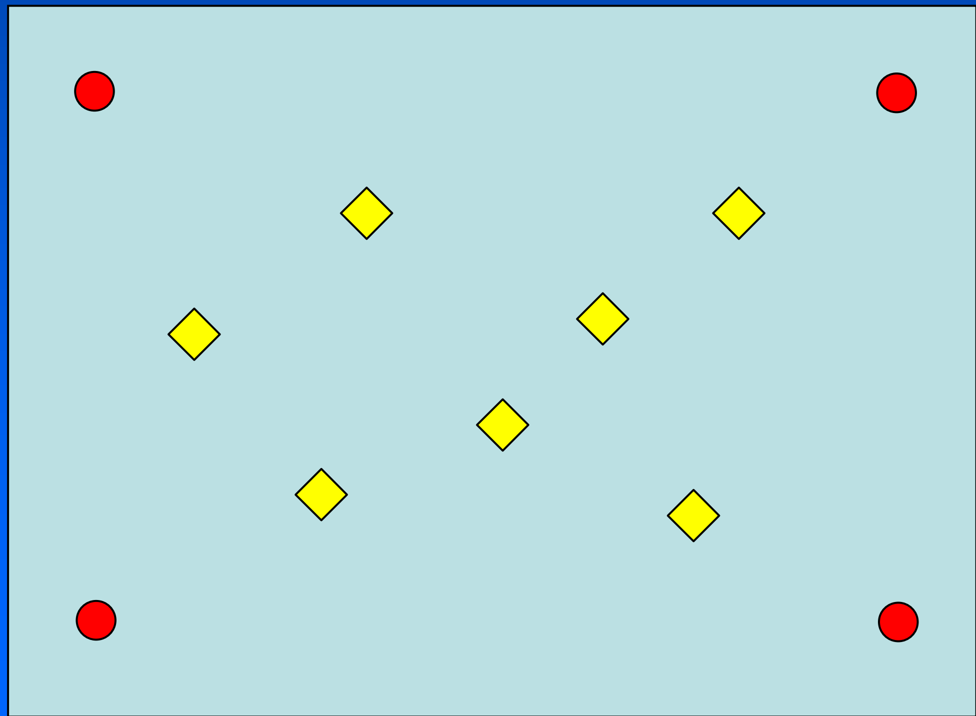
$$e_S = f_M\left(p_1,\ldots,p_n,\bar{\theta}_1,\ldots\bar{\theta}_k\right) - f_S\left(p_1,\ldots,p_n,\theta_1,\ldots\theta_k,\Delta x,\Delta t,\ldots\right)$$

Calibration with Coarse Models

- Build response surface for difference between ‘fine’ and ‘coarse’ models

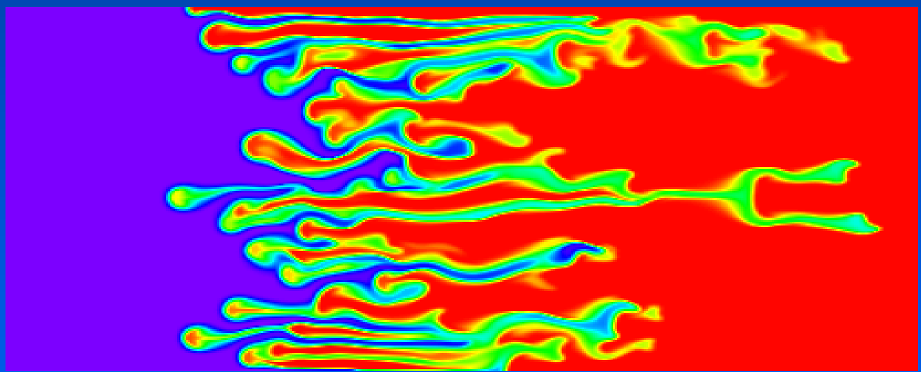
– calculated error

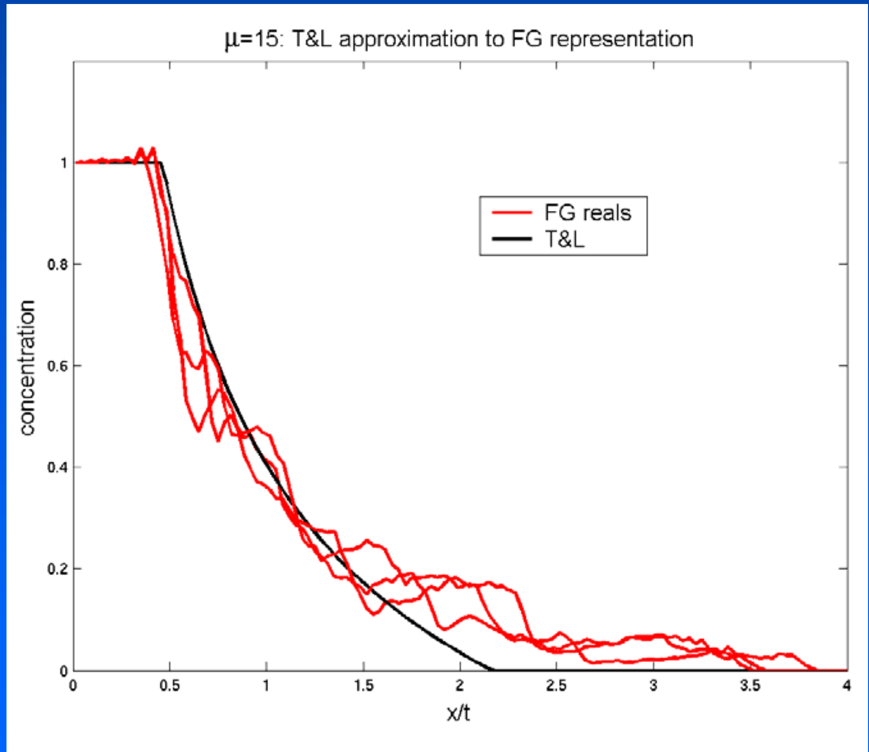
– interpolated error



One Parameter Example

- See Los Alamos Science, vol 29

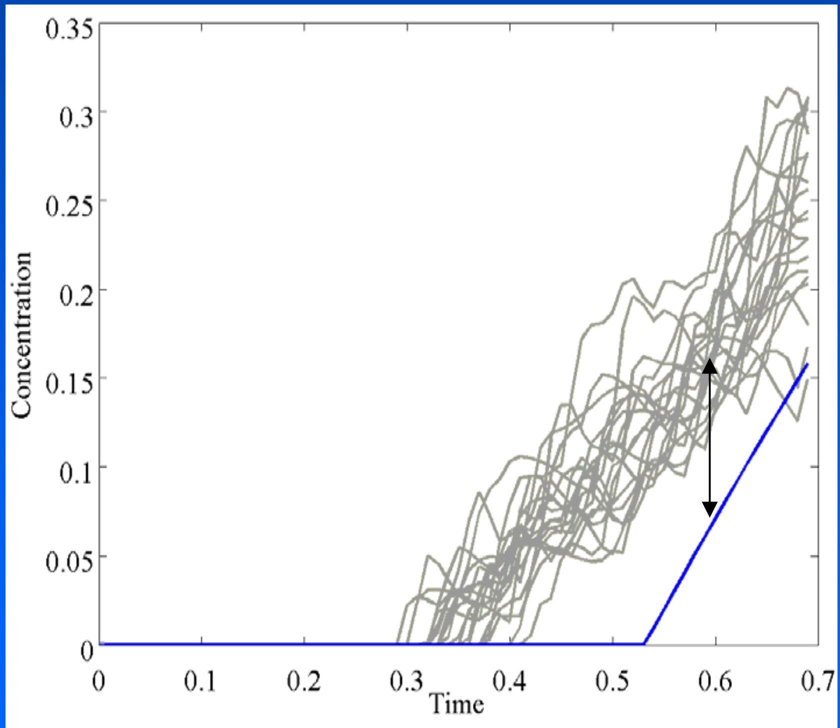


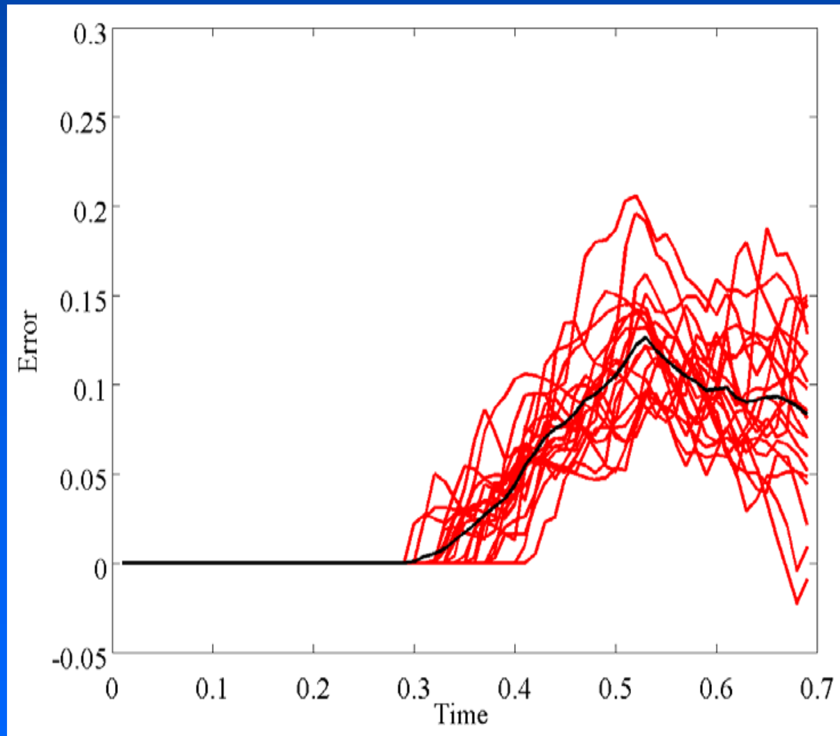


Calculate Time Varying Solution Errors

$$e_i = FG_i - CG_i$$

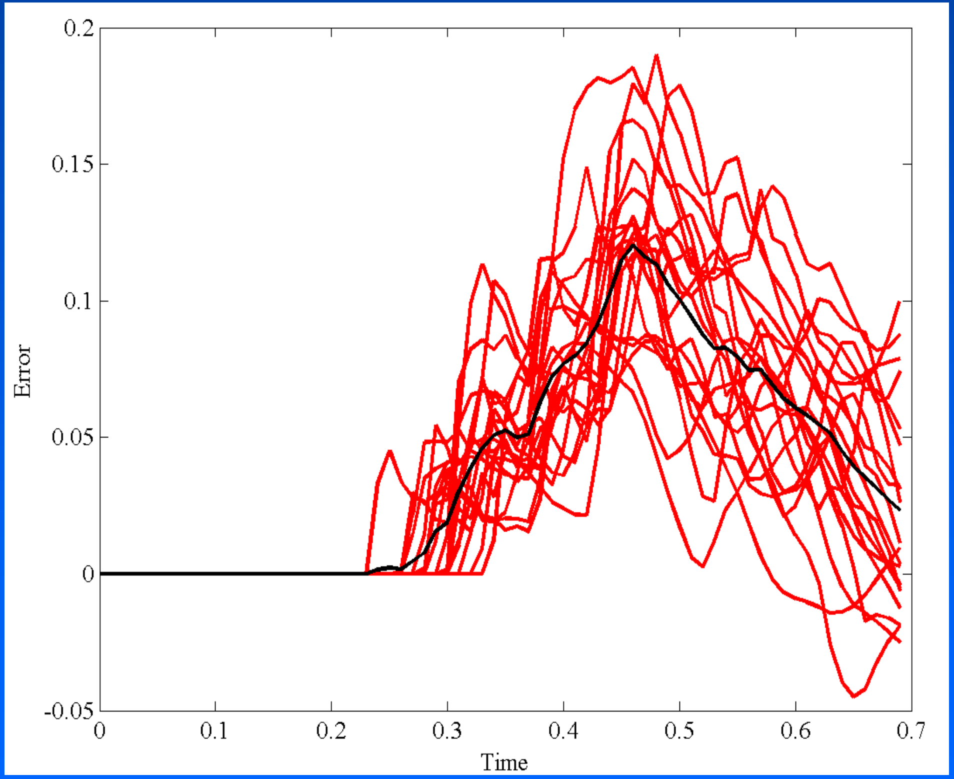
$$\bar{e} = \frac{1}{n}\sum_{i=1}^{n} e_i$$

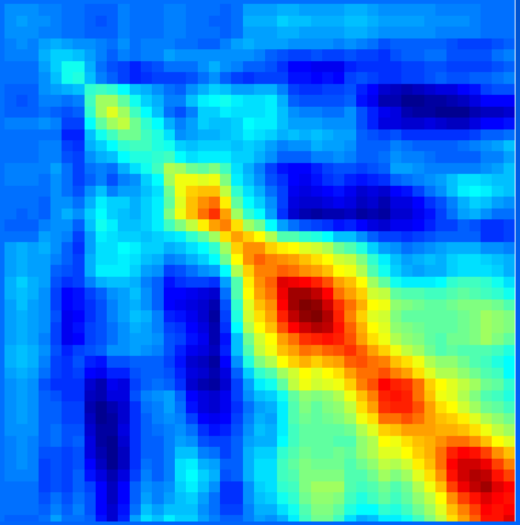




Determine Covariance Structure

$$C_{se}(s,t) = \frac{1}{n-1}\sum_{j=1}^{n}\left(e_j(t)-\overline{e}(t)\right)\left(e_j(s)-\overline{e}(s)\right)$$



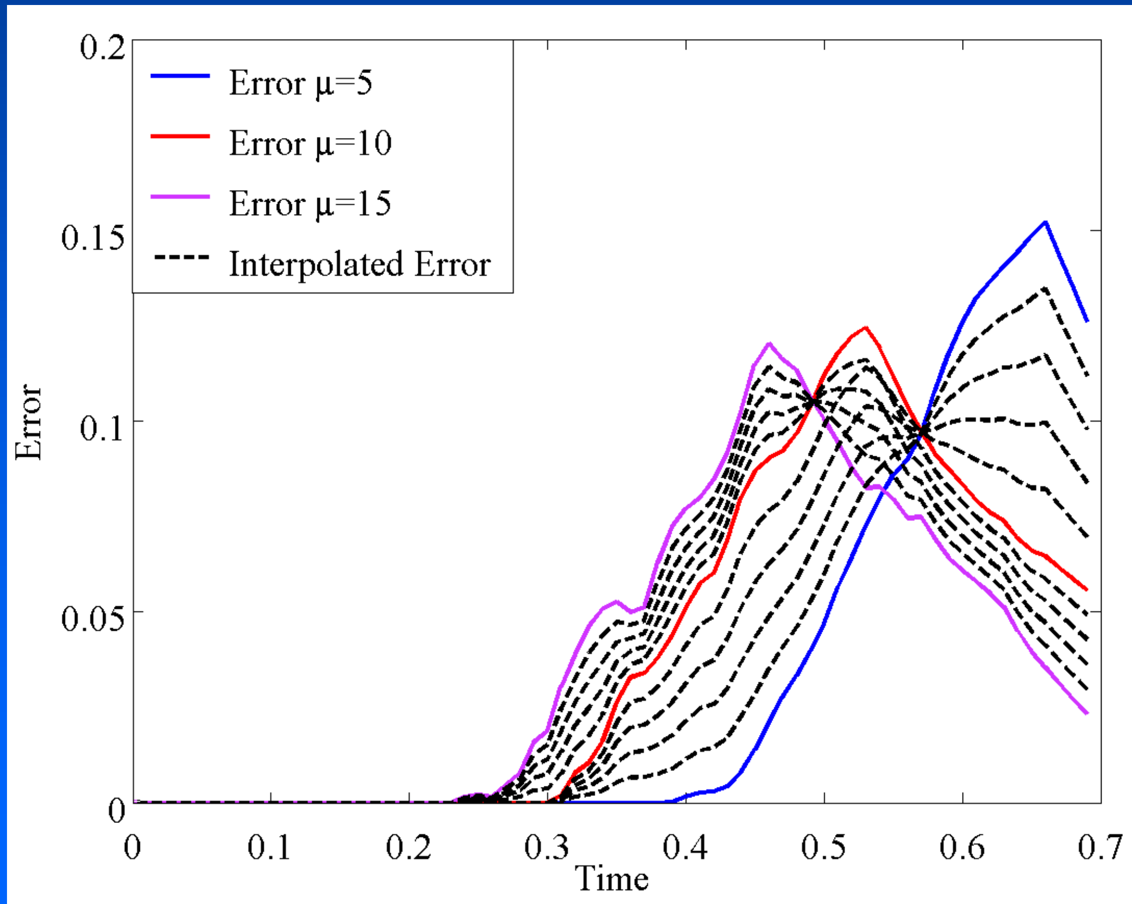


Interpolate for Mean Error and Covariance

Example:

Linear interpolation for mean error and covariance

More sophisticated scheme yields minor improvements

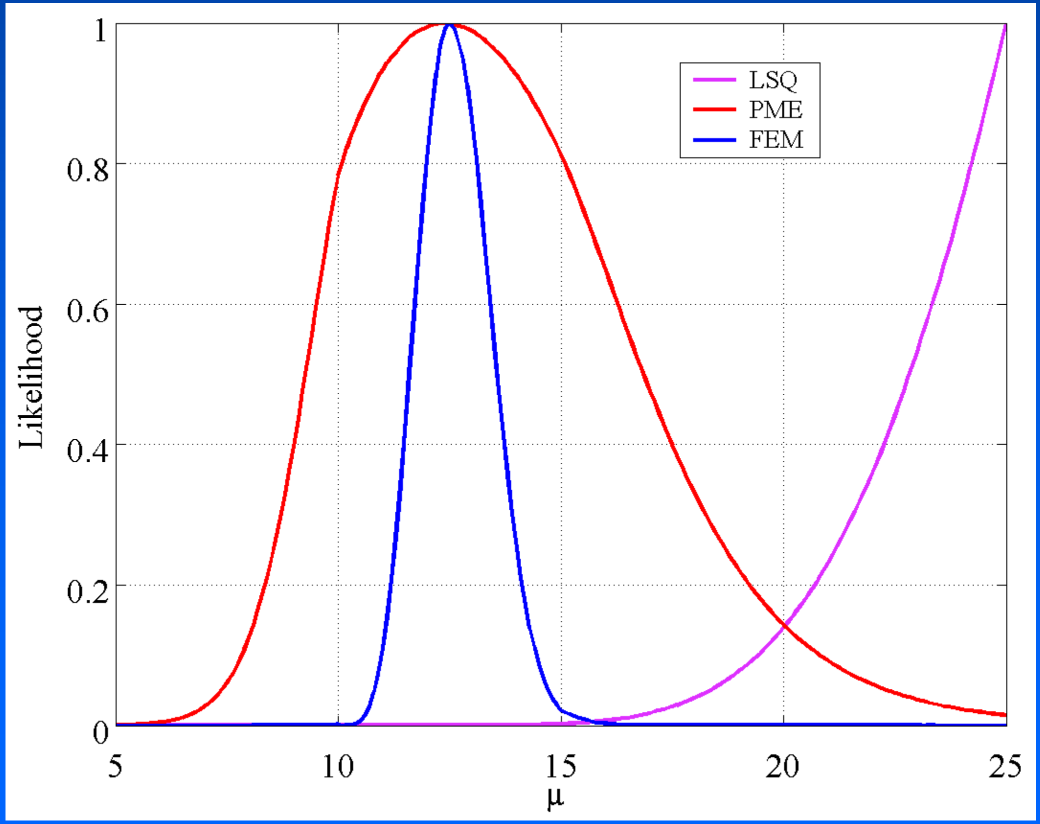


Results

True oil viscosity, μ = 13

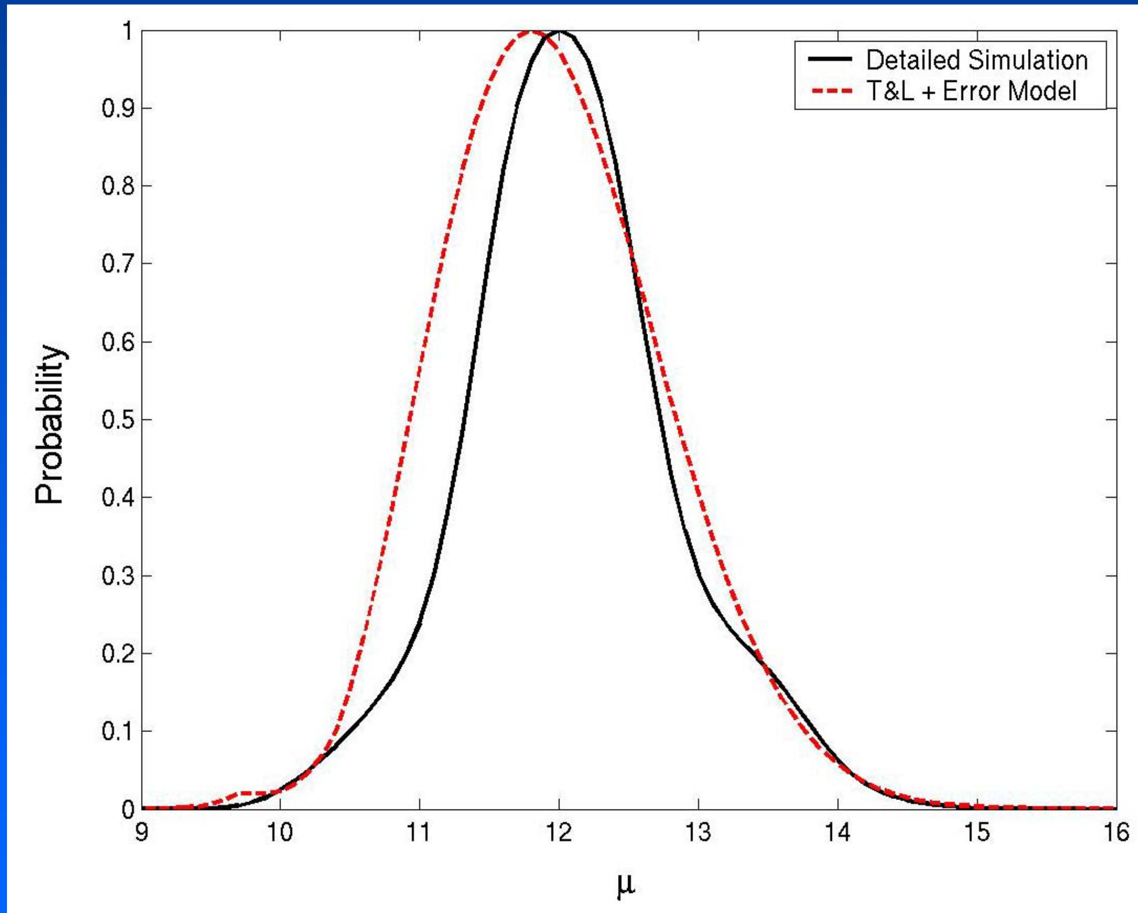
Maximum likelihood, σ=12.5

Effect of bias significantly reduced

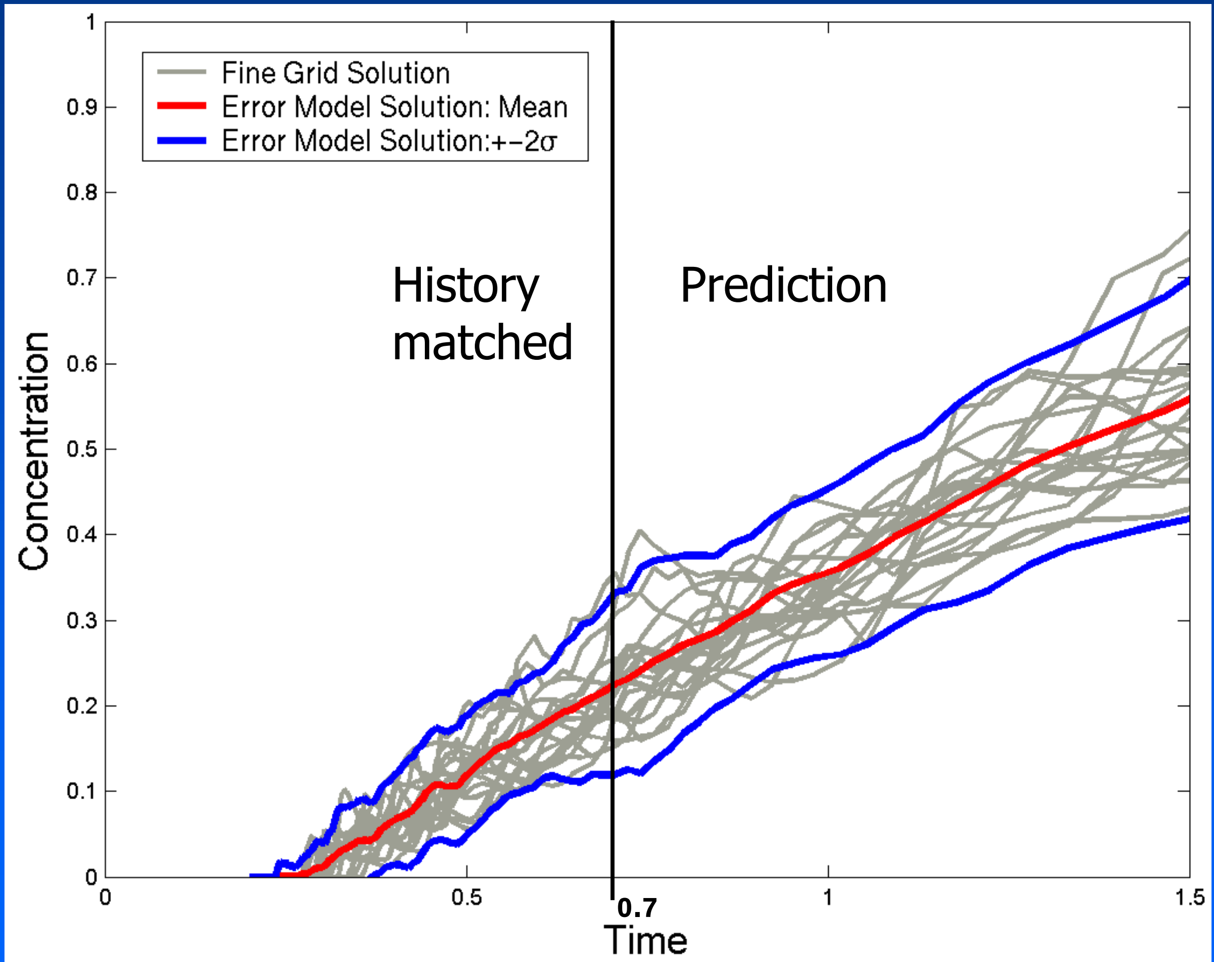


How Accurate is the Error Model?

Likelihood curves



Prediction


Fine Grid Solution
Error Model Solution: Mean
Error Model Solution:+−2σ
History matched
Prediction
Concentration
Time
0.7
0
0.5
1
1.5
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9

Alternative Interpolation Approach

- Based on Kennedy & O'Hagan (2001)
  - Error depends on unknown parameters and independent variables in specific way
  - Discrepancy function only of independent variables
  - Described in *J. Roy. Stat. Soc. B*, 2001

Kennedy and O'Hagan Approach

$$z_i = \boxed{\rho\eta(x_i, \theta) + \delta(x_i)} + e_i$$

z_i observations

ρ scalar

η simulator

x_i variable input

θ history match input

δ model inadequacy

e_i observation error

$$\varsigma(x_i) = \rho\eta(x_i, \theta) + \delta(x_i)$$

truth

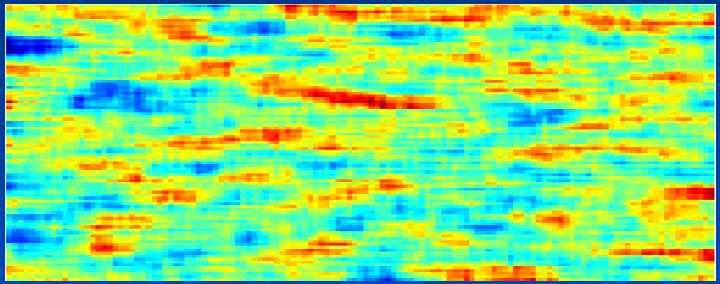
Errors for Varying Grid Sizes

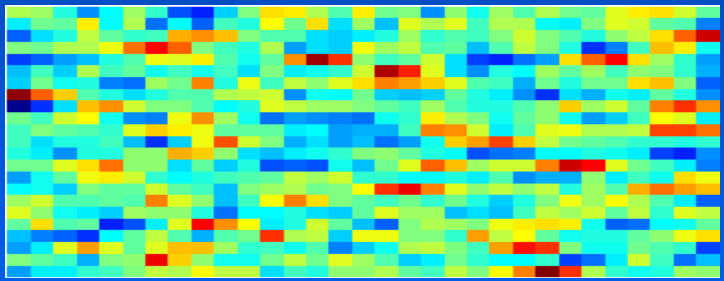
- 128x96=12288 cells (fine)

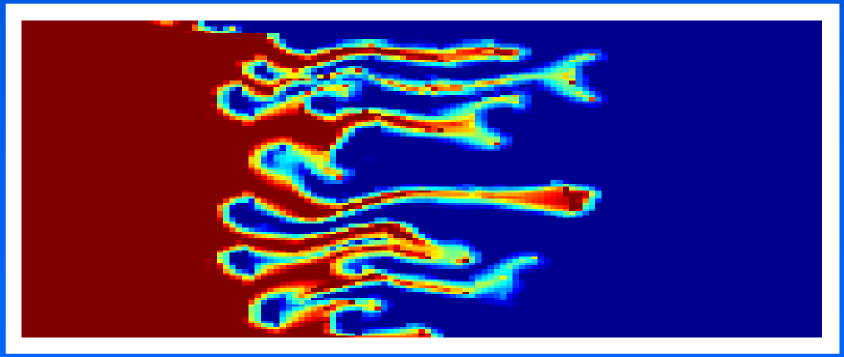
  64x48 = 3072 cells

  32x24 = 768 cells

  16x12 = 192 cells



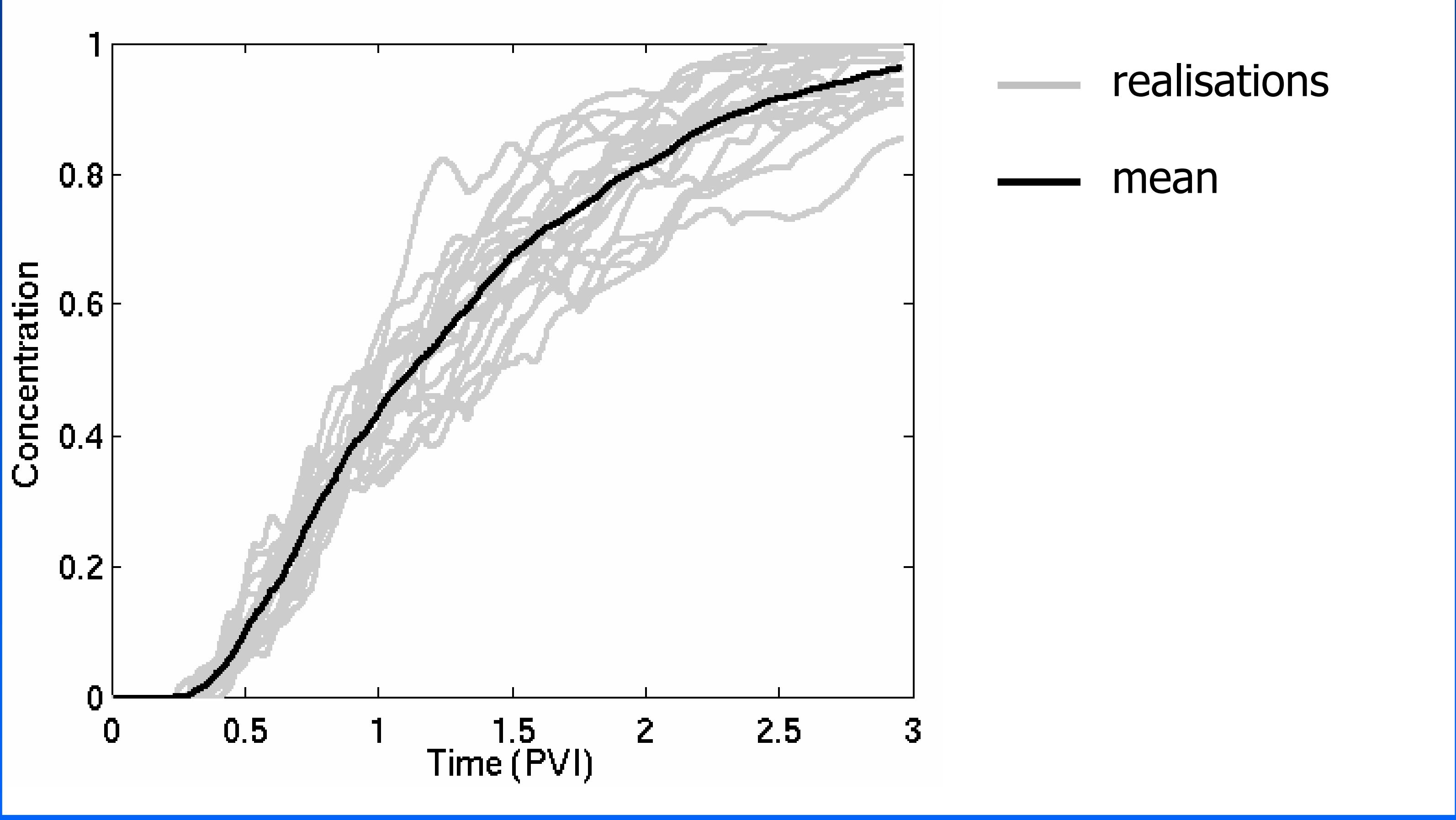




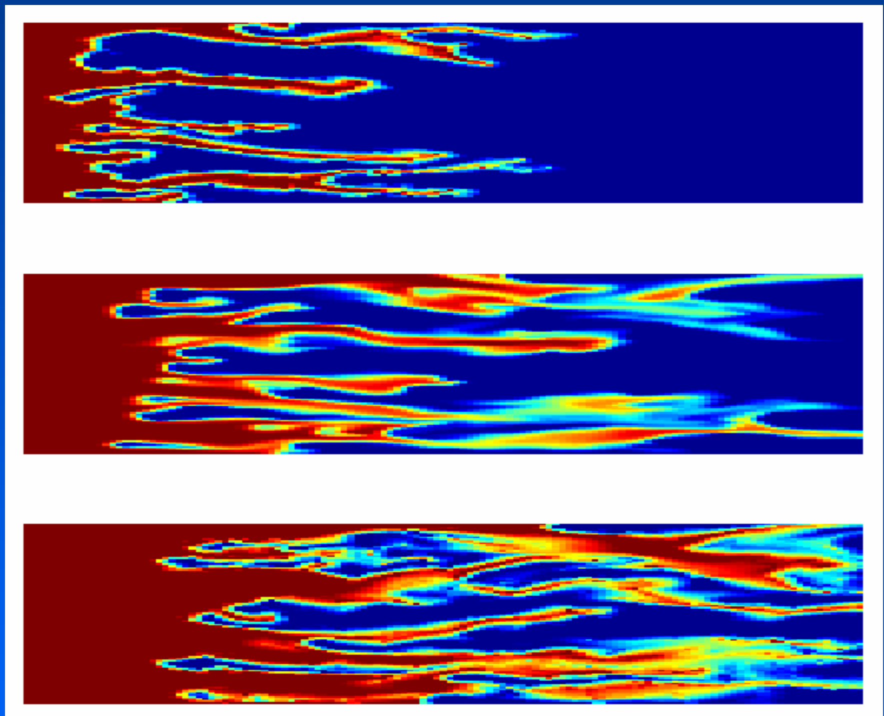
Perm:

- Mean 1
- Var in log 0. 5
- Cx 0.2 Cy 0.05

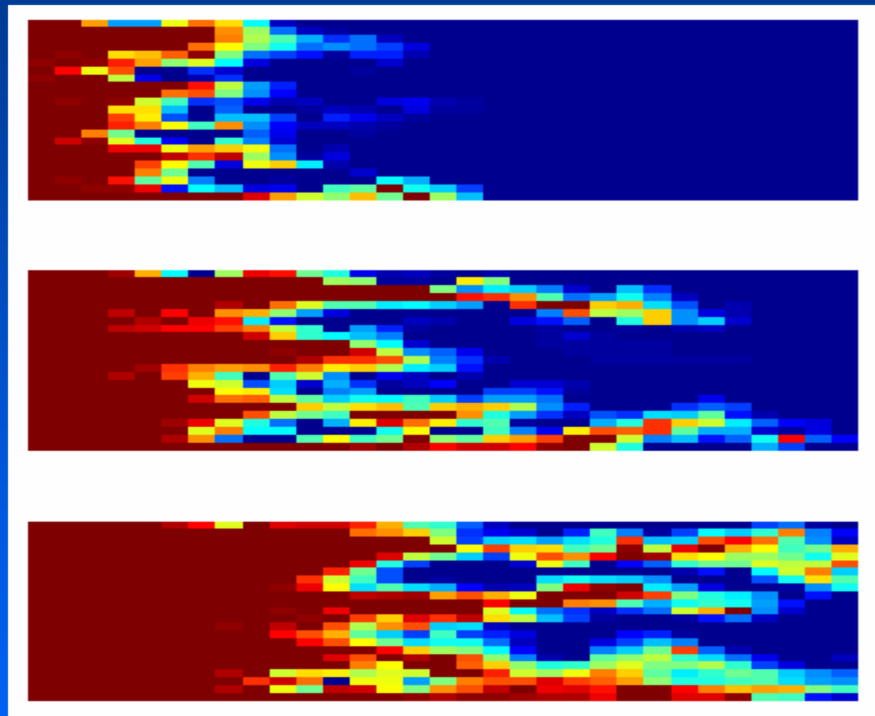
Variability in Solution



Saturation Plots

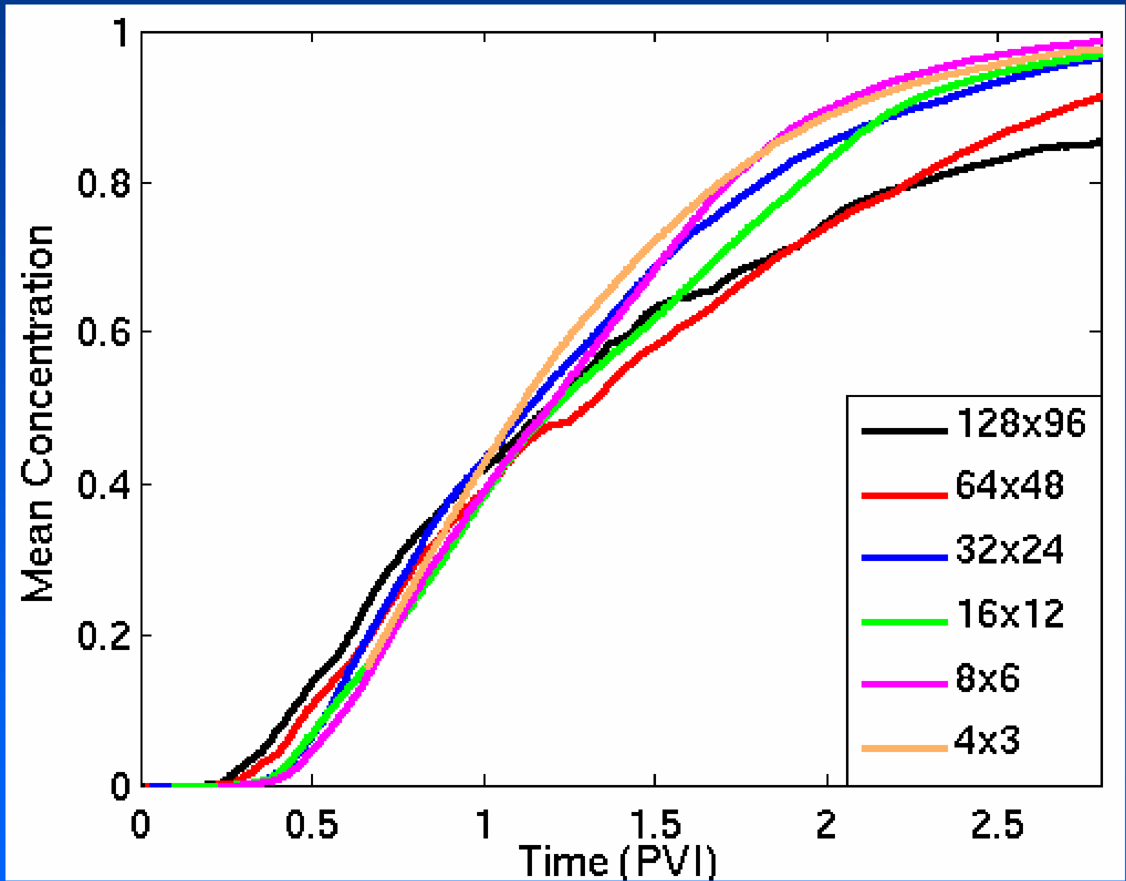


Grid 128x96

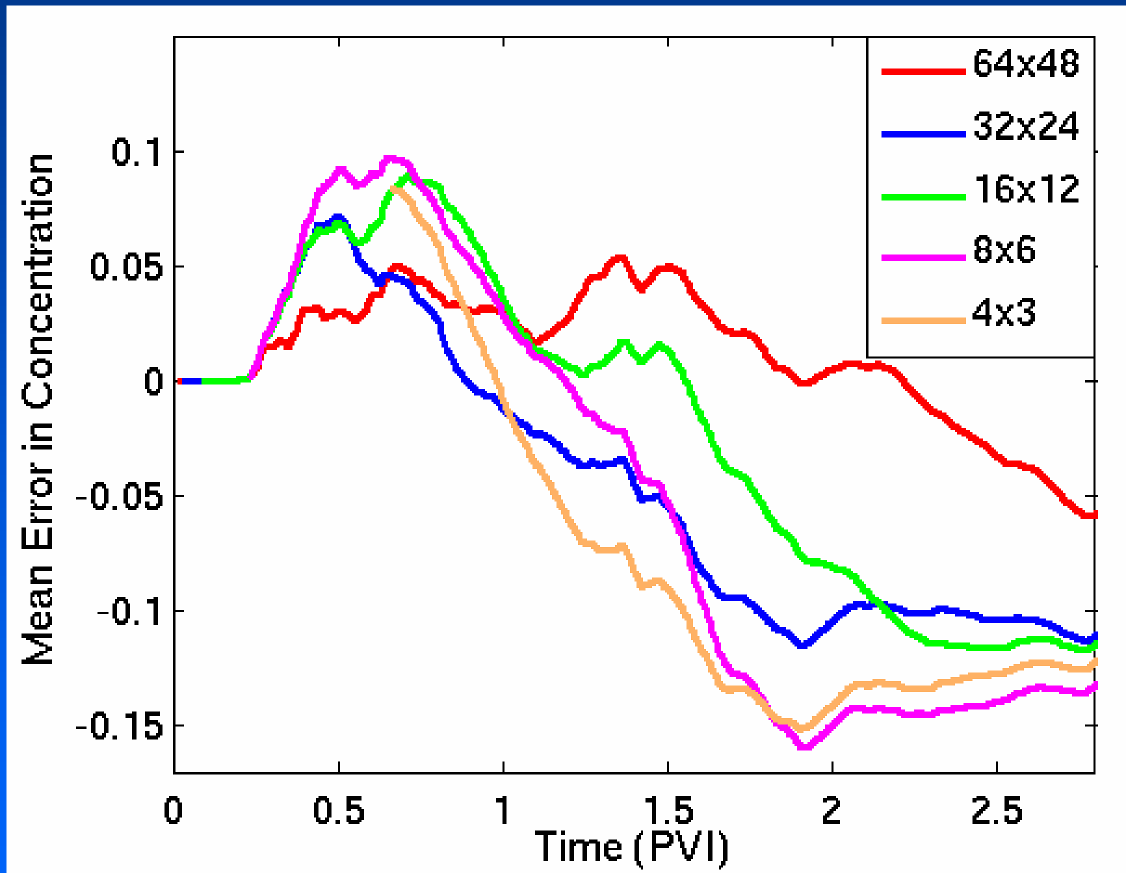


Grid 32x24

Solution as a Function of Grid Size



Error as a Function of Grid Size



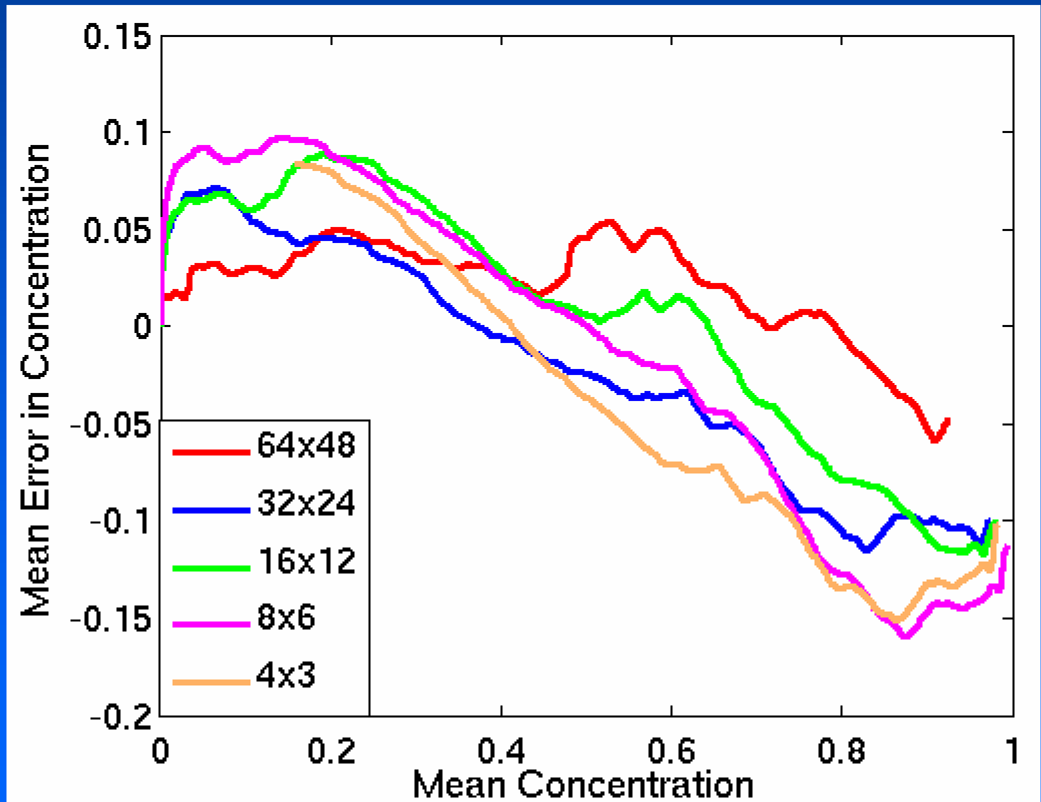
$$E = FG - CG$$

Fine grid is 128x96

Error as a Function of Solution

$$e(x_i,\theta) = (\rho - 1)\eta(x_i,\theta) + \delta(x_i)$$

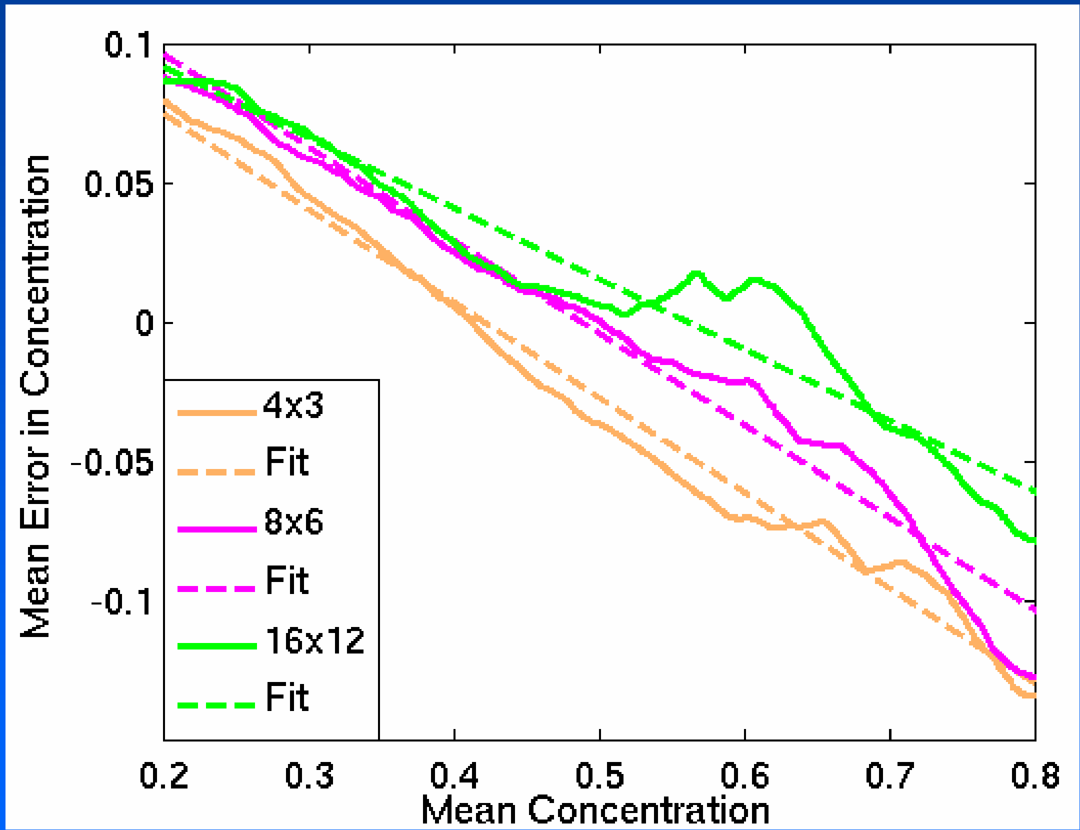
$e(x_i,\theta)$



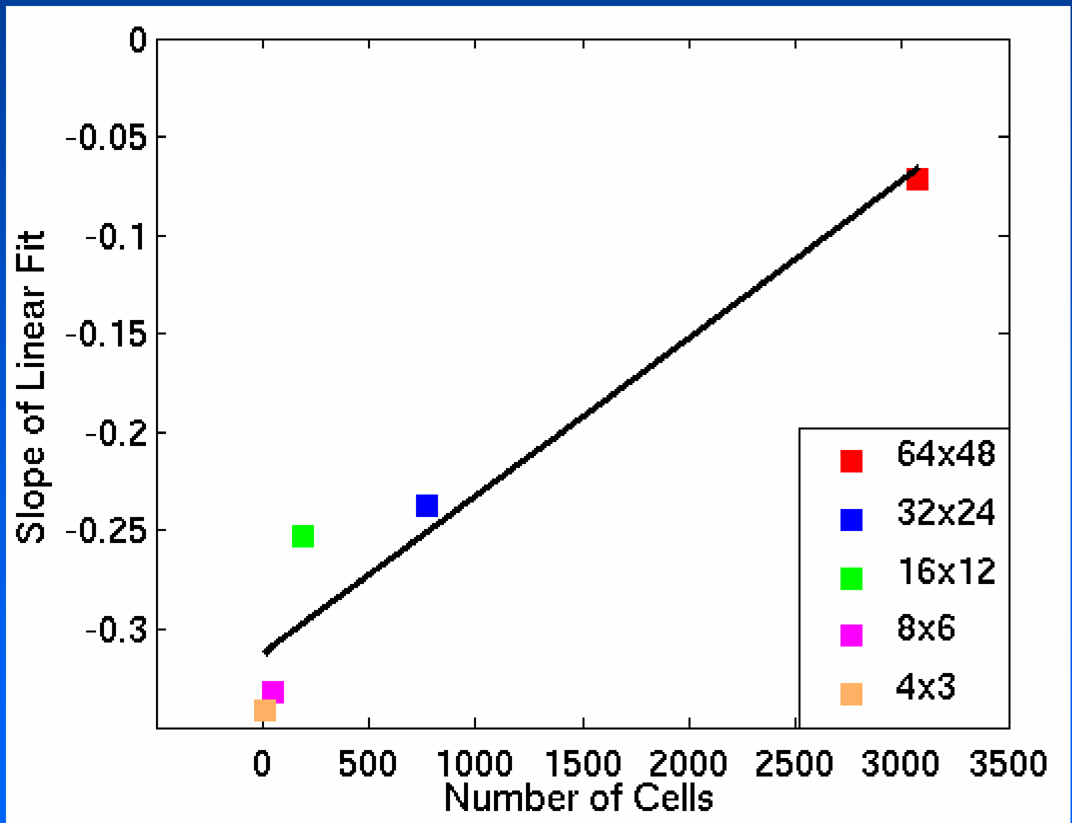
$\eta(x_i,\theta)$

θ fixed

Determining Parameterized Fit

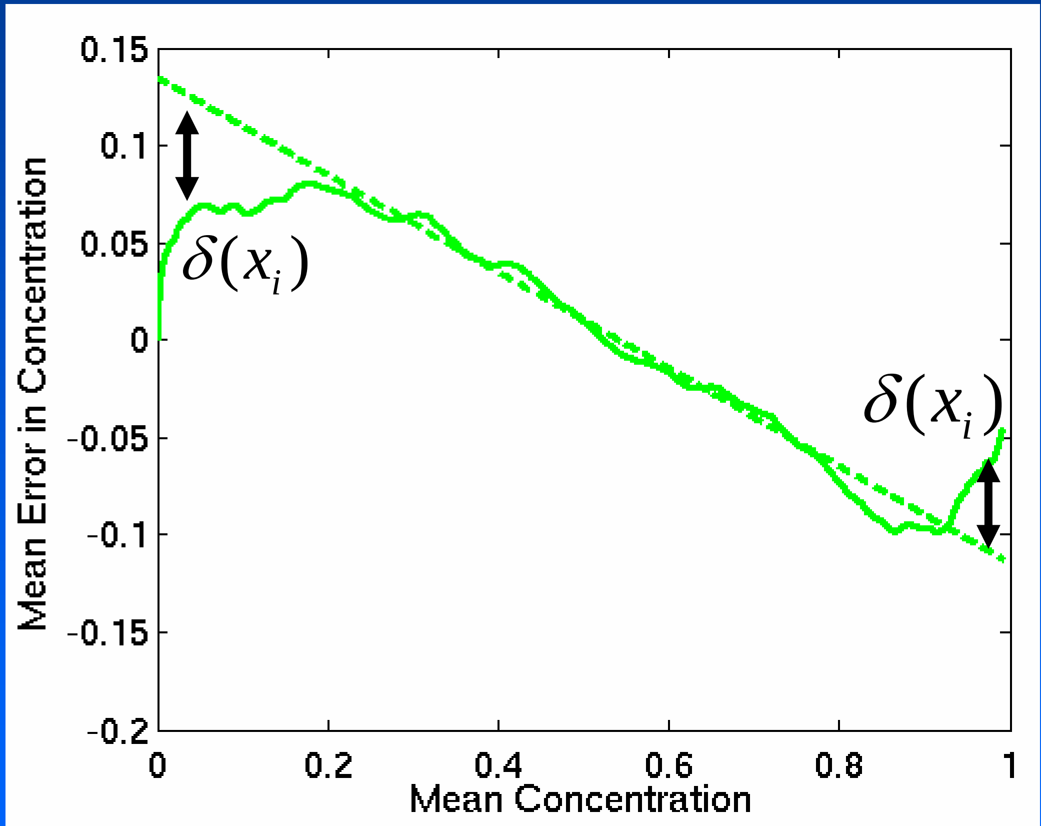


Determining Parameterized Fit: Slope



Viscosity constant

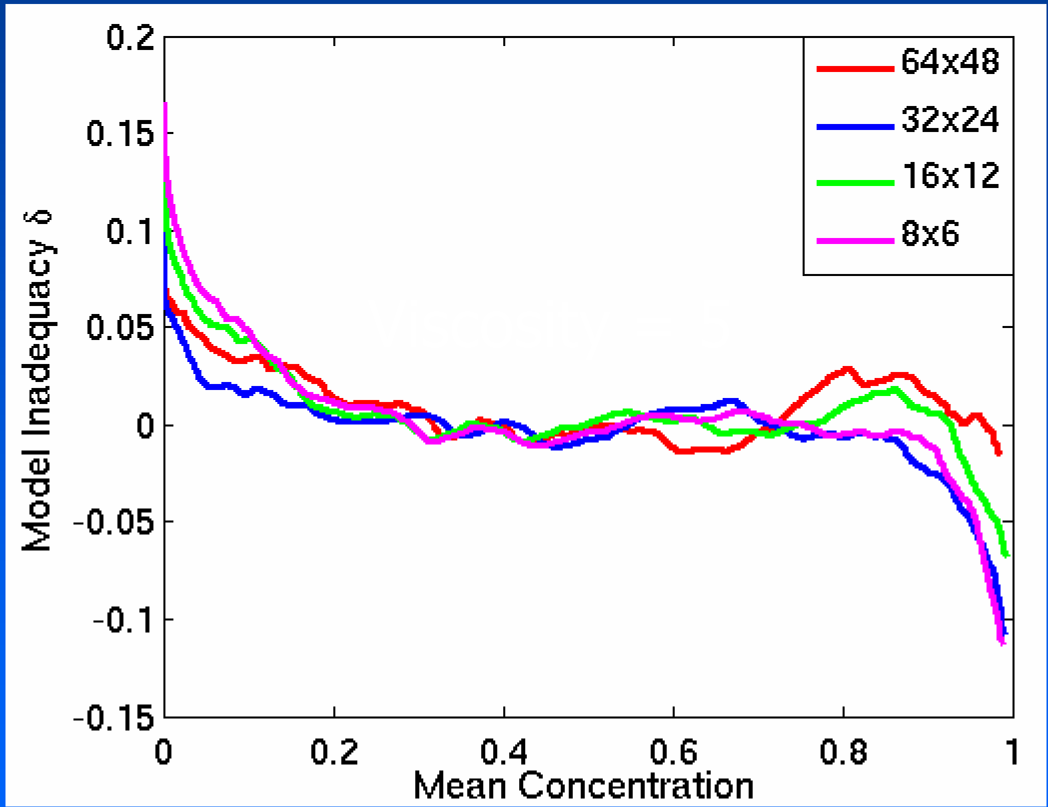
Determining Model Inadequacy



Viscosity = 5

Grid 16x12

Model Inadequacy with Grid Size



Viscosity = 5

Error Modelling for the Lorenz Equations

- Lorenz equations are a simple model for chaotic convective flows in the atmosphere
- Lorenz equations

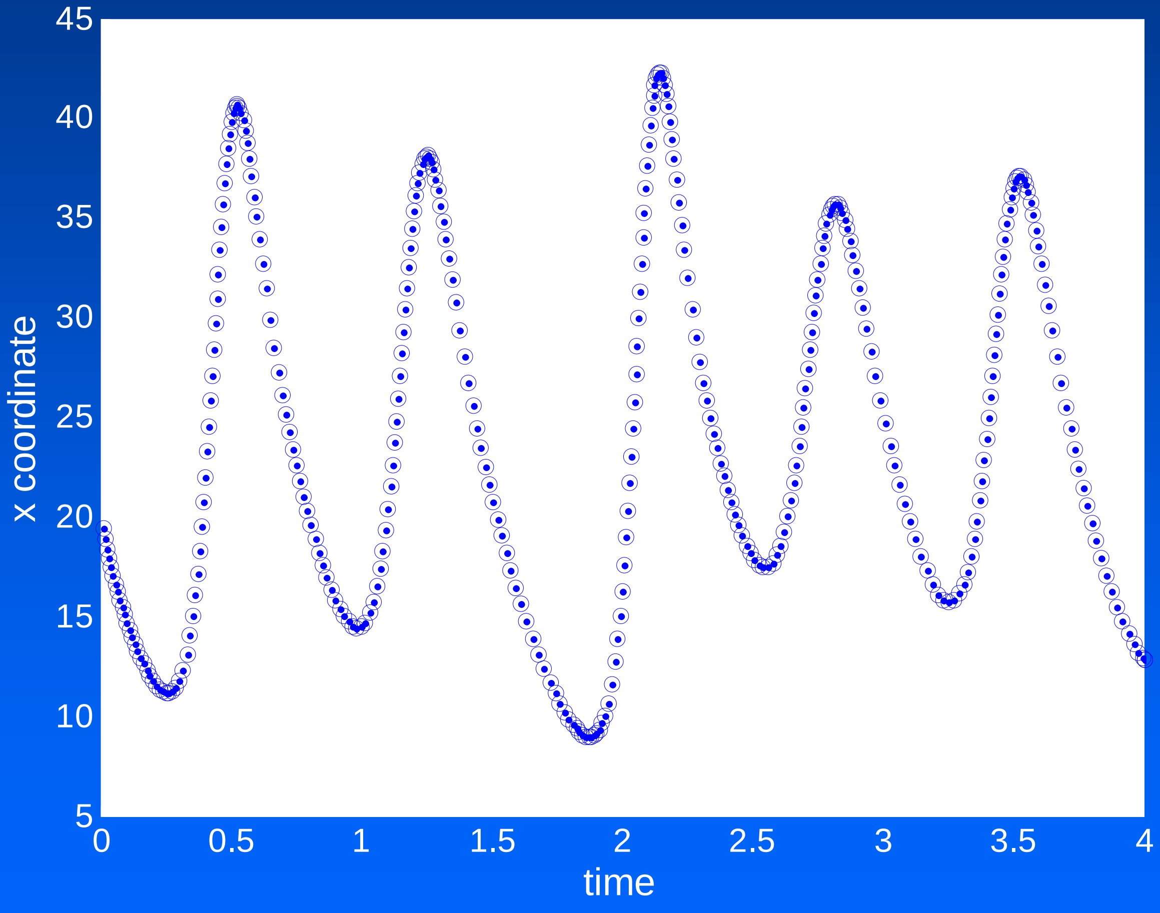
$$\frac{dx}{dt} = \sigma(y - x) \qquad \sigma = 10$$

$$\frac{dy}{dt} = x(\rho - z) - y \qquad \rho = 28$$

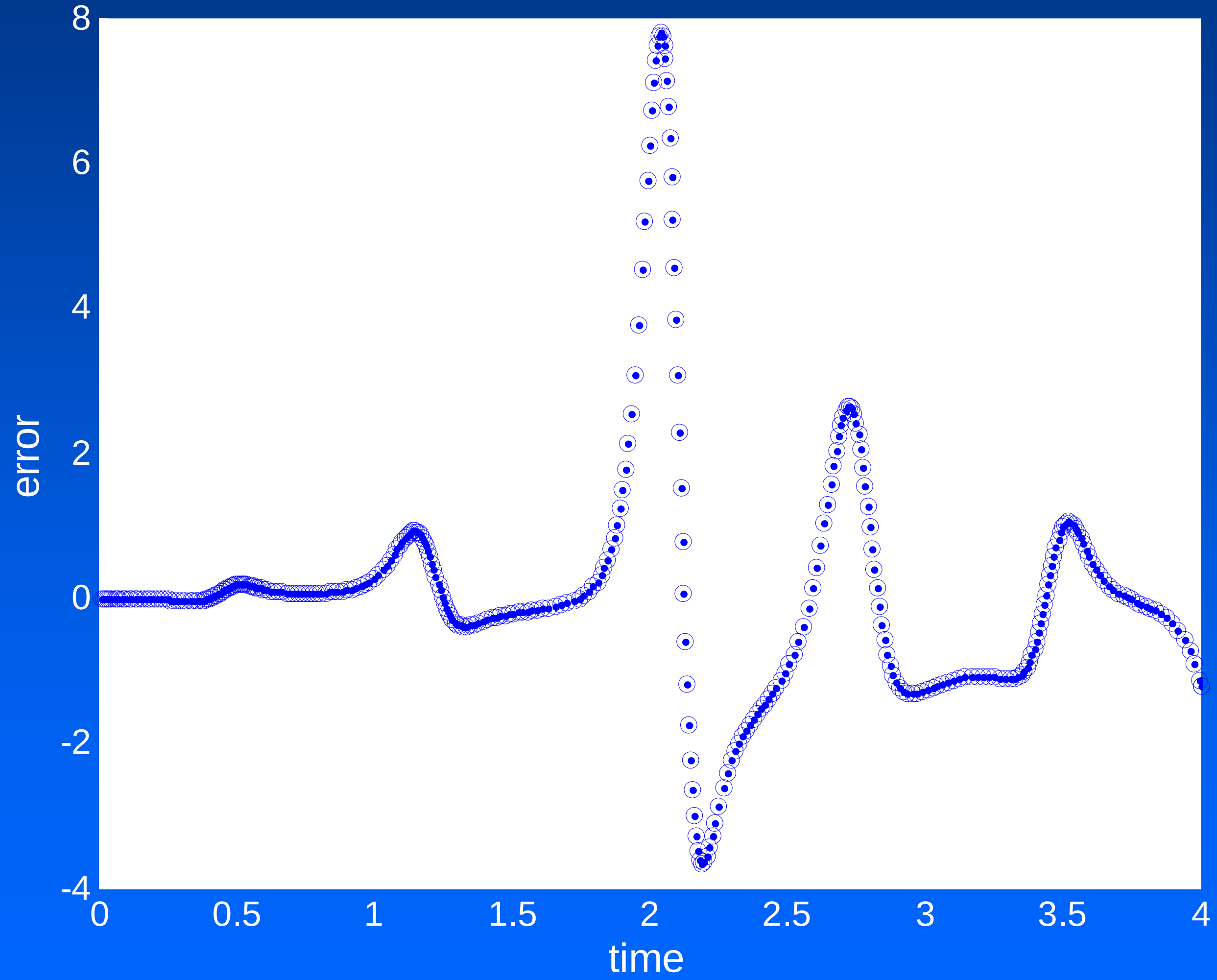
$$\frac{dx}{dt} = xy - \beta z \qquad \beta = 8/3$$

- Model error
  - Use wrong value of ρ = 28.1
  - Match for σ and β

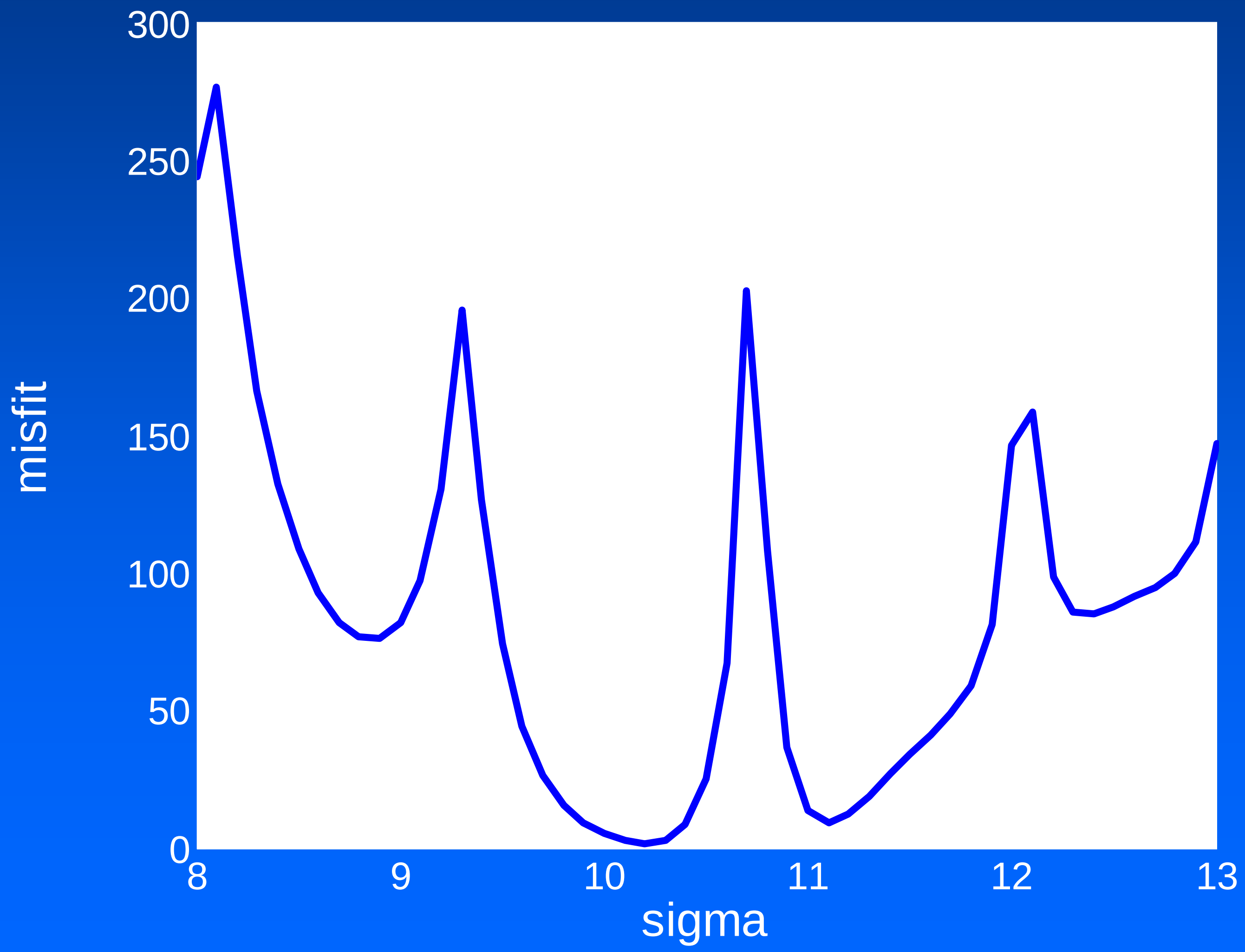
Lorenz Equations Time Series – Exact Model



Error: Exact - Wrong Model

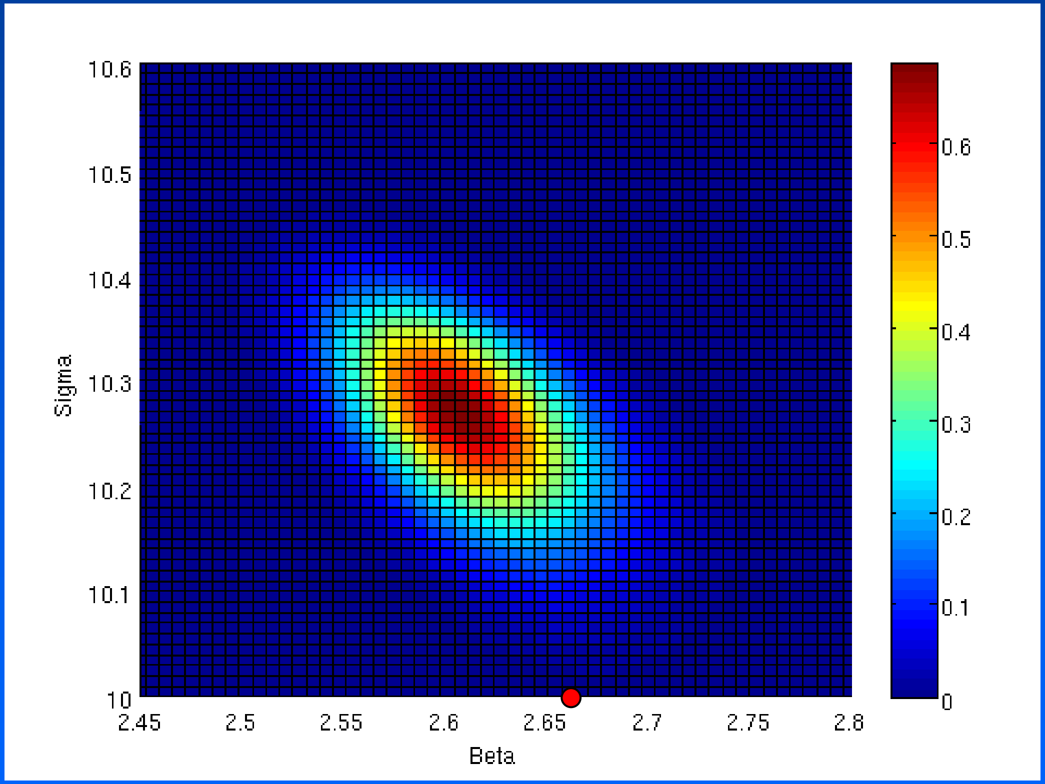


Misfit Cross Section for Wrong Model

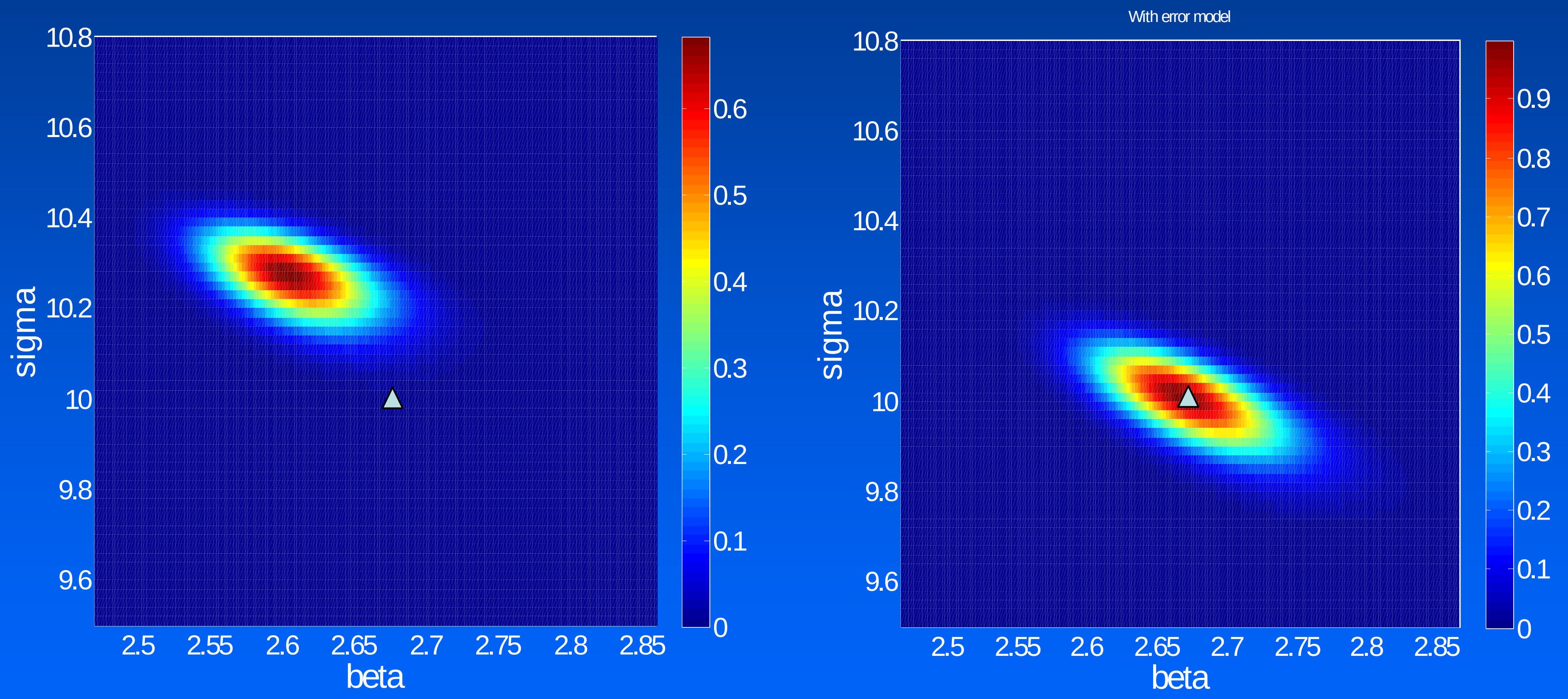


Likelihood cross section – wrong model

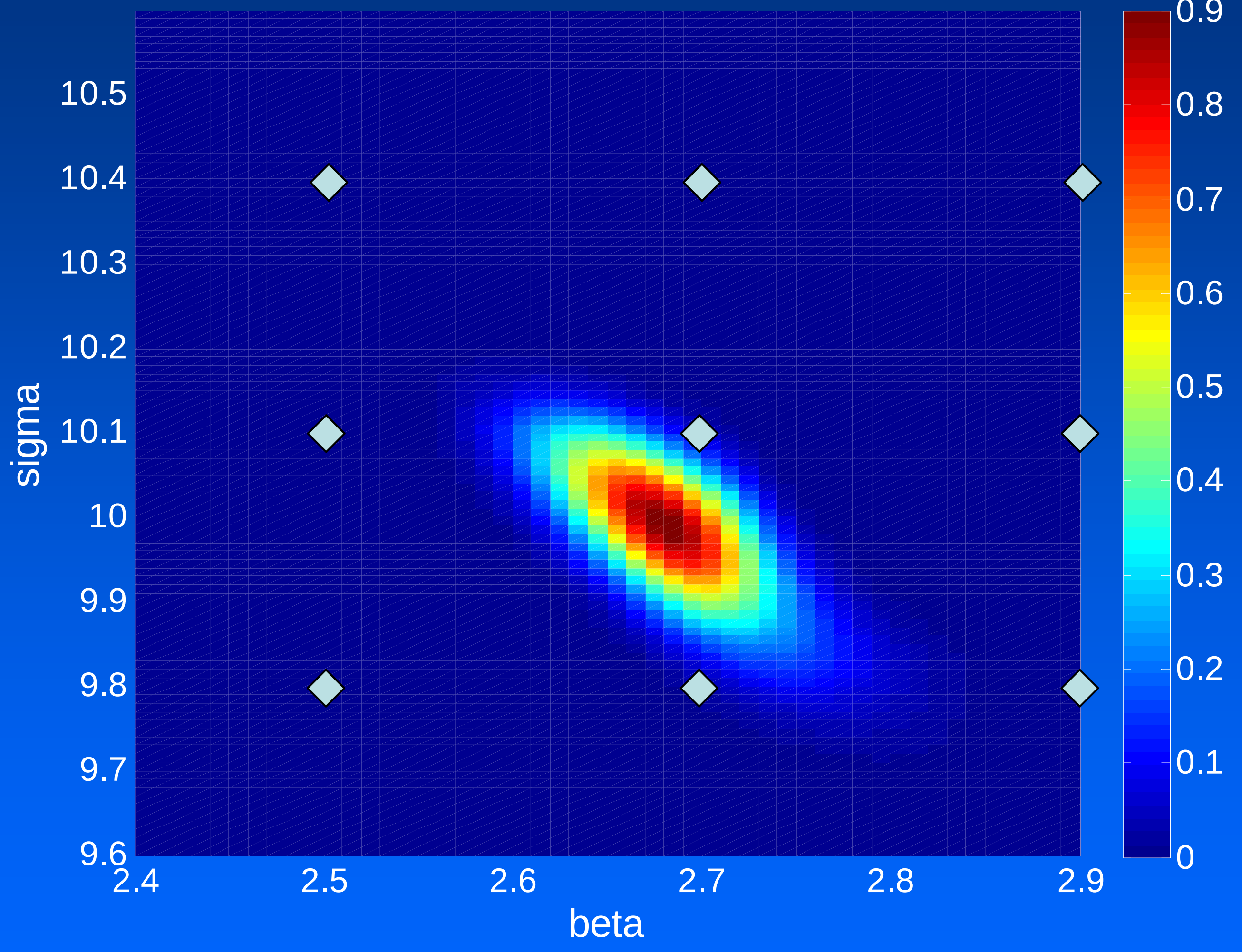
- True values: $\beta = 8/3, \sigma = 10$



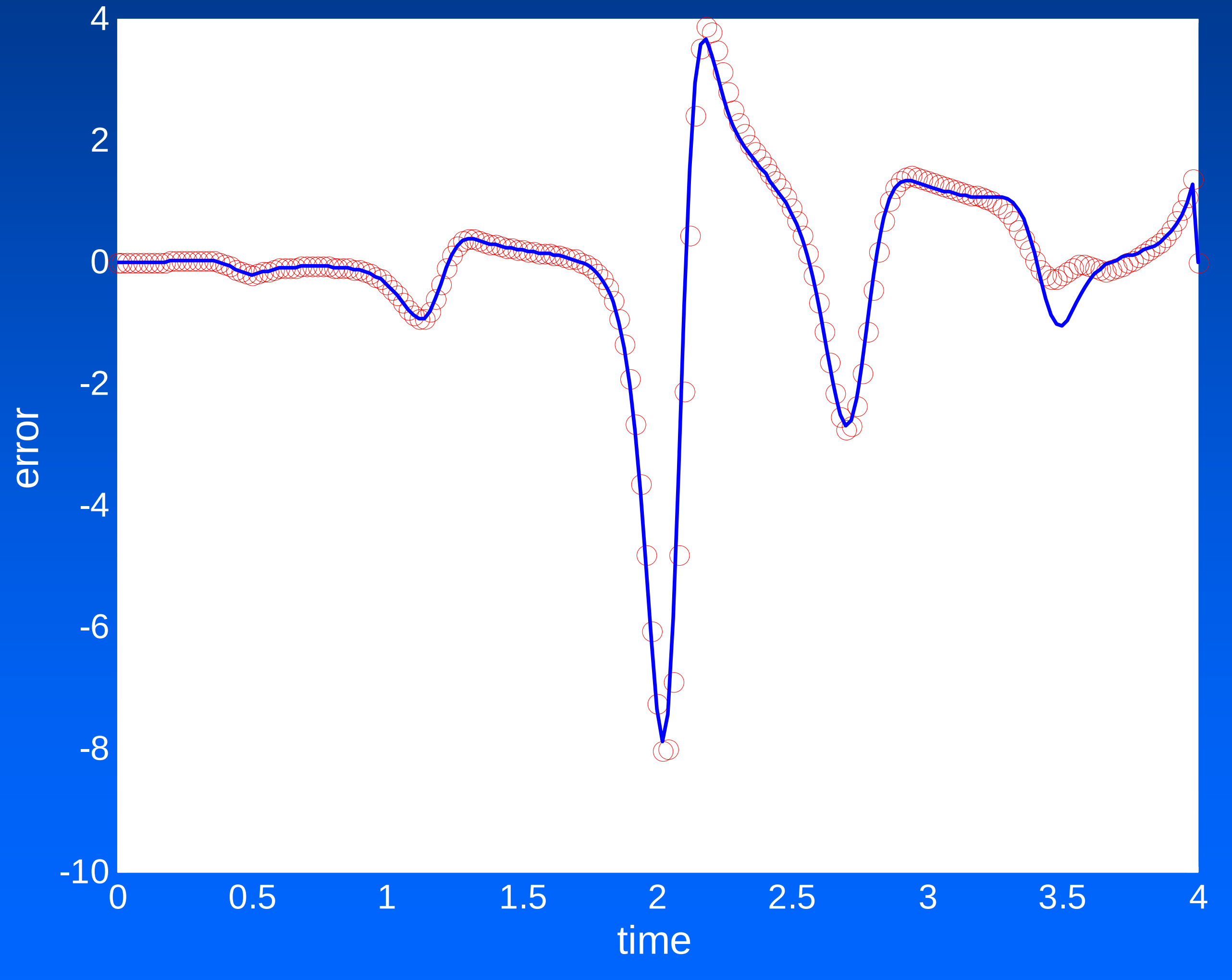
Likelihood Plot with and without Error Model



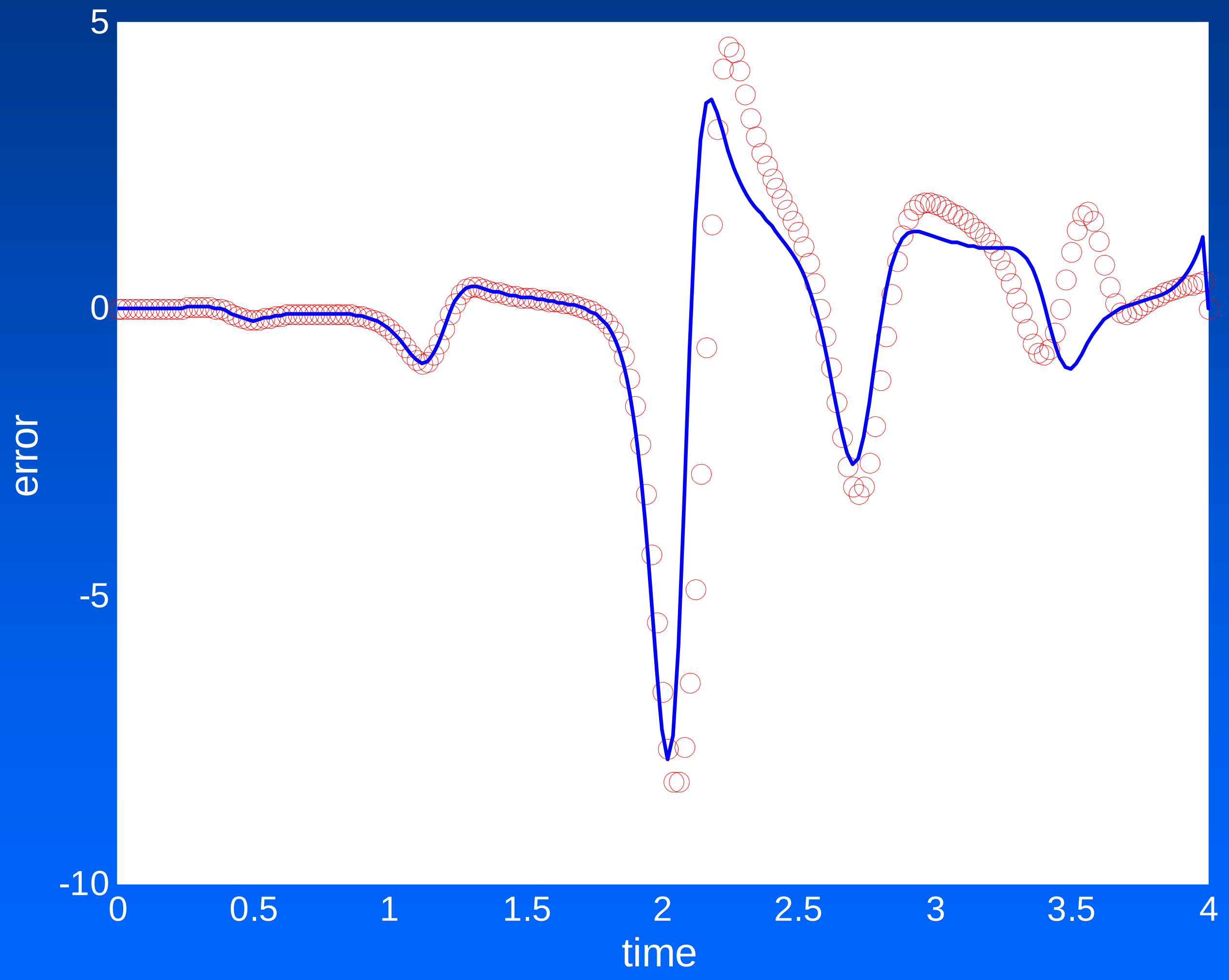
Error Model – 3 x 3 Base Points



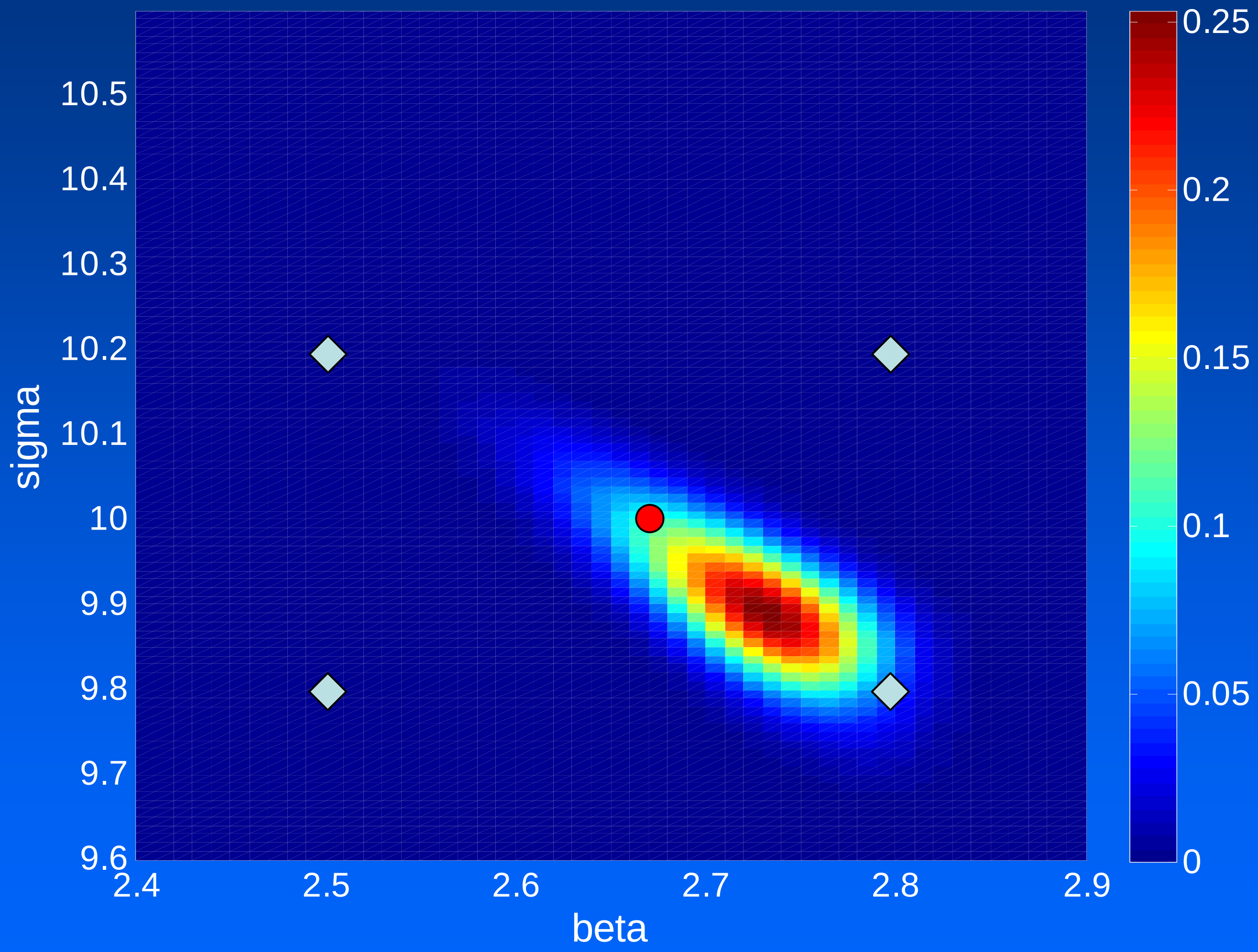
Interpolated Errors – 3 x 3 grid



Interpolated Errors – 2 x 2 grid



Error Model – 2 x 2 Base Points



Summary

- Solution Error Models
  - Compute errors due to known effects
    - Sub-grid physics
    - Inadequate resolution
  - Function of unknown parameters and independent variables
  - Practical issue
    - Need to compute error model with only limited number of solves of fine model